



UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK

-----X
JANELL WINFIELD, TRACEY STEWART,
and SHAUNA NOEL

Plaintiffs,

-against-

CITY OF NEW YORK,

Defendant.
-----X

KATHARINE H. PARKER, UNITED STATES MAGISTRATE JUDGE

INTRODUCTION

As has been widely reported in the news, New York City's current goal is to build or preserve 200,000 below-market-rate apartments by 2024 as part of its affordable housing programs. See William Neuman, *De Blasio Says City Will Hit Affordable-Housing Goal 2 Years Early*, N.Y. TIMES, (Oct. 24, 2017), <https://www.nytimes.com/2017/10/24/nyregion/de-blasio-affordable-housing-goal-2-years-early.html>. This case involves a challenge to a component of New York City's affordable housing programs—the Community Preference Policy. New York City residents who want affordable housing and otherwise meet the qualifications (including the income requirements) can apply for certain affordable housing units through a lottery system. There are far more people seeking affordable housing than there are units available. The Community Preference Policy at issue applies to some, but not all affordable housing lotteries.

The Policy sets aside 50% of the affordable housing units to be distributed through the lottery for people who live in the “community district”¹ in which the housing is located.²

Plaintiffs, three African-American residents of New York City, seek affordable housing and have applied for housing through lotteries, but have not yet been selected to be interviewed for affordable housing developments. They claim that the Community Preference Policy has a disparate impact on African-American and Latino applicants in “neighborhoods of opportunity,” which they assert are predominantly white.³ They also claim that the Community Preference Policy perpetuates racial segregation in the City and that its promulgation and application constitutes intentional discrimination in violation of the federal Fair Housing Act, 42 U.S.C. §§ 3604, *et seq.*, and the New York City Human Rights Law, N.Y.C. Admin. Code §§ 8-107, *et seq.*

The issues presently before the Court relate to the City’s production of electronic documents. The genesis of the parties’ dispute was a disagreement over what search terms should be used in connection with the City’s review of electronic documents collected from five

¹ New York City is broken down in 59 community districts.

² In or about 2002, the administration of Mayor Michael Bloomberg increased the percentage of units for which preference would be given to residents of the community district in which the development was being built from 30% to 50%. (Doc. No. 16 (“FAC”) ¶ 83.) Plaintiffs allege that the change was implemented without evaluating whether the increase would perpetrate residential segregation. (See FAC ¶¶ 85, 154-56.)

³ For context, the Court notes that according to 2010 census data, New York City is a “majority minority” city, meaning that its population is approximately 34% White, 23% Black, 29% Hispanic, and 13% Asian. See N.Y.C. Dep’t of City Planning, *NYC2010: Results From The 2010 Census*, available at: <https://www1.nyc.gov/assets/planning/download/pdf/data-maps/nyc-population/census2010/pgrhc.pdf> (last visited Nov. 27, 2017). The 2010 census data also reflects that White residents represents the majority population in 24% of the City’s census tracts. See The Furman Center for Real Estate & Urban Policy, *The Changing Racial And Ethnic Makeup Of New York City Neighborhoods*, available at: http://furmancenter.org/files/sotc/The_Changing_Racial_and_Ethnic_Makeup_of_New_York_City_Neighborhoods_11.pdf (last visited Nov. 27, 2017).

custodians from the Department of City Planning (“DCP”) and Steven Banks, the Commissioner of the City Human Resources Administration/Department of Social Services (“HRA”) and head of the Department of Homeless Services (“DHS”). Since then, Plaintiffs have complained of various alleged deficiencies in the City’s electronic document review process, including (i) the scope of the City’s review of documents gathered from DCP and Banks, and (ii) the City’s alleged over-designation of documents in the prior review populations as non-responsive, which Plaintiffs claim has affected the reliability of the City’s predictive coding processes as a whole. In order to increase transparency into the City’s document review process, Plaintiffs seek, among other things, an order directing the City to provide Plaintiffs with samples of non-privileged documents collected from the Department of Housing Preservation & Development (“HPD”), the Mayor’s Office, DCP, and Banks that the City has designated as “non-responsive” in its review.

Having now reviewed the City’s *in camera* submissions, as well as sample documents provided by Plaintiffs, **Plaintiffs’ application is granted in part and denied in part as discussed below.**

BACKGROUND

A. Case Summary

HPD is the City agency primarily responsible for the creation and preservation of affordable housing in the City. It promotes development of new affordable housing, as well as the preservation and renovation of existing housing that contains affordable housing units. In many cases, the City helps finance the development projects through, among other things,

loans, tax credits, and exemptions. HPD also oversees the allocation of certain affordable housing units to eligible applicants through a lottery. The Community Preference Policy applies to only some of the lotteries and affordable housing units that become available in the City.

In addition to the housing promoted by HPD, certain housing developments in New York City have been incentivized by tax exemptions provided under Section 421-a of New York's Real Property Tax Law. The Community Preference Policy also has been applied to some of these so called "421-a developments." Pursuant to a policy announced by the City in 2016, 50% of the community preference units in certain 421-a developments have been set aside for homeless residents of the City whose shelter or last-known address is in the community district where the 421-a development is built.

DCP is the City's primary land use agency and it is responsible for the City's Mandatory Inclusionary Housing program, a zoning initiative developed by HPD and DCP that requires certain types of residential projects to include affordable housing. DCP also oversees the Uniform Land Use Review Procedure ("ULURP"), which applies to many affordable housing projects to the extent they require zoning or other land use actions.⁴ DCP is the lead agency in formulating neighborhood plans and coordinating infrastructure investments, as well community support services. DCP also coordinates the City's consolidated plans and

⁴ Under the ULURP procedure, DCP receives certain applications for housing developments, after which the Community Board for the community district in which the development is proposed holds a public hearing and makes a recommendation about the development. The Borough President and Borough Board also make a recommendation as to the development. DCP then holds a public hearing and either approves, modifies, or disapproves the application. If the development project is approved, City Council then may review the application and hold a public hearing because it approves certain tax exemptions and approves budget allocations for affordable housing. After the City Council's review, the Mayor may review and veto any action taken by the City Council. The City Council can override a veto by a 2/3 vote.

certifications regarding its federal obligations to affirmatively further fair housing pursuant to a rule promulgated by the U.S. Department of Housing and Urban Development.

Plaintiffs charge in their Complaint that the City has been and continues to be characterized by extensive residential segregation on the basis of race, ethnicity, and national origin in all 59 community districts. (FAC ¶¶ 1, 47-48.) They assert that segregation can be traced to historical restrictions and intentional discrimination by all categories of actors in the housing market, including government entities, developers, landlords, and others. (FAC ¶¶ 2-3.) They claim the Community Preference Policy has a disparate impact on African-Americans and Latinos insofar as they do not have an equal opportunity to compete for affordable housing in all neighborhoods, particularly those that are “higher opportunity.” (*See, e.g.*, FAC ¶ 102.) They also contend that the continuing application of the Community Preference Policy amounts to intentional discrimination based on race because the City’s “decisions to establish, maintain, and expand the policy: (a) were made in the face of a history of discrimination and segregation encouraged by and participated in by the City; (b) were made knowing, or being deliberately indifferent to, the policy’s clear disparate impact on opportunity to participate on equal terms and its tendency to perpetuate segregation; (c) constituted choices to reject more pro-integrative alternatives; (d) are reflective of the City’s consciousness of what policies it thought that particular racial and ethnic groups “wanted,” as well as other race-awareness; and (e) responded to racially- and ethnically-influenced community and political opposition.” (FAC ¶ 8; *see also* FAC ¶¶ 9, 113-37, 171.) Plaintiffs also allege that the City’s failure to comply with its

obligations under federal law to affirmatively further fair housing is evidence of its intent to discriminate.

Plaintiffs assert that the Community Preference Policy is popular among many community boards, local politicians, and advocacy groups “based on a desire to preserve existing racial or ethnic demographics or culture of a neighborhood or community district.” (FAC ¶¶ 158, 160.) Plaintiffs further assert that “the City [has] allowed itself to be influenced by such race- or ethnicity-based positions or by fear that an abandonment of the [Community Preference Policy] would generate race- or ethnicity-based opposition from these community boards, local politicians, and advocacy groups.” (FAC ¶ 161.)

The City denies the allegations of discrimination and states that it is deeply committed to fair housing and the creation of affordable housing. Former HPD Commissioner Vicki Been has explained that the Community Preference Policy “is intended to ensure that local residents, many of whom have deep roots in the community and have persevered through years of unfavorable living conditions, are able to remain in their neighborhoods as those neighborhoods are revitalized” in part through development. (Doc. No. 18 ¶ 8.) She has stated that neighborhoods “often resist approving land use actions required to allow greater density or site affordable housing because of concern about the other types of burdens that development may impose” such as noise and danger from construction, crowding and strain on existing infrastructure, and the like. (Doc. No. 18 ¶ 8.) She has further stated that the Community Preference Policy “ensures that neighborhoods see that new growth and investment in affordable housing provide some benefits to local residents to offset those burdens” and “makes

it possible for the City to overcome” neighborhood resistance to development and achieve its affordable housing goals. (Doc. No. 18 ¶ 8.)

B. *The Motion To Dismiss*

At the outset of this case, the City moved to dismiss the Complaint for lack of standing and failure to state a claim. The Honorable Laura Taylor Swain denied the City’s motion to dismiss on October 24, 2016, finding, among other things, that the Plaintiffs had (1) “adequately alleged that they are persons who ‘will be injured by a discriminatory housing practice that is about to occur,’ 42 U.S.C. § 3602(i), and have standing to challenge the [Community Preference Policy],” (2) “satisfied their burden of demonstrating plausibly that the Community Preference Policy perpetuates segregation” in support of their disparate impact discrimination claim, and (3) pled sufficient facts to plausibly support their disparate treatment claim. *Winfield v. City of New York*, No. 15-cv-5236 (LTS) (DCF), 2016 WL 6208564, at *5-7 (S.D.N.Y. Oct. 24, 2016).

In the decision, Judge Swain noted that “[d]iscovery may well prove that removal of the Community Preference Policy would have no impact on the racial and ethnic makeup of eligible affordable housing applicants (and ultimate recipients of housing) whether local residents within a community district are given a preference or not.” *Id.* at *6. The Court also noted that certain facts pled, such as a history of discriminatory zoning and housing policies, concentrating public housing in minority neighborhoods, and racial steering in housing projects, supported the plausibility of Plaintiffs’ claims. *Id.* at *7.

C. *Discovery*

Plaintiffs have sought wide-ranging discovery, which the City has resisted vigorously. This Court has issued various rulings consistent with Federal Rule of Civil Procedure 26(b)(1) to ensure that discovery is focused on information that is relevant and proportional to the needs of this case, considering “the importance of the issues at stake in the action, the amount in controversy, the parties’ relative access to relevant information, the parties’ resources, the importance of the discovery in resolving the issues, and whether the burden or expense of the proposed discovery outweighs its likely benefit.” Fed. R. Civ. P. 26(b)(1). This included directing phased discovery and limiting some of Plaintiffs’ discovery demands.

The first phase of discovery has focused on production of data from the affordable housing lotteries needed for disparate impact analysis, depositions of key City witnesses, and production of documents concerning the Community Preference Policy and the reasons for its promulgation and amendments, as well as documents concerning the City’s defenses, including community resistance to affordable housing projects.

Plaintiffs have sought, among other things, documents related to affordable housing policies more generally, City policies related to zoning or homelessness, interactions with the U.S. Department of Housing and Urban Development (“HUD”), and strategies for mitigating displacement of individuals from gentrifying communities. During a conference on February 16, 2017, the Court narrowed a number of Plaintiffs’ discovery requests, including by topical scope and the number of custodians that the City needed to search in response to 23 of Plaintiffs’ document requests.

The parties also disagreed as to the custodians from whom documents would be collected. Ultimately, after meet and confers and rulings by this Court, the City was ordered to collect documents, including electronic documents, from 50 custodians. These included 44 individuals who worked within HPD and the Mayor's Office, 5 individuals who worked in DCP, and Banks.

Pursuant to the Court's directives, the City began its document review process with HPD and the Mayor's Office. As it pertained to ESI, the parties disagreed as to the search terms that would be applied to the electronic documents collected from these custodians. Ultimately, the City applied a set of search terms that was heavily negotiated by the parties, with some guidance from the Court, to the custodians from HPD and the Mayor's Office. The City then began reviewing the electronic documents from these custodians.

Plaintiffs lodged numerous complaints about the pace of discovery and document review, which initially involved only manual linear review of documents. Accordingly, given the volume of documents collected, this Court directed the City to complete linear review as to certain custodians and begin using Technology Assisted Review ("TAR") software (also commonly referred to as "predictive coding") to hasten the identification, review, and production of documents responsive to Plaintiffs' document requests.

TAR allows parties to prioritize and/or categorize documents for purposes of document review and has been shown to produce more accurate results than manual review. *See Rio Tinto PLC v. Vale S.A.*, No. 14-cv-3042 (RMB) (AJP), 2015 WL 872294, at *3 (S.D.N.Y. Mar. 2, 2015); *Da Silva Moore v. Publicis Groupe*, 287 F.R.D. 182 (S.D.N.Y. 2012) (noting that linear

manual review was too expensive when there were over three million emails to review and that computer-assisted review can save both parties significant time and money in document review and is at least as, if not more, accurate as manual review). For TAR to work properly, the producing party must prepare a training, or seed set, of responsive and non-responsive documents to train the computer system how to distinguish between them. For purposes of this review, the City trained its TAR system using documents that it had already reviewed during its linear review, as well as by completing several training rounds of additional documents that had not yet been reviewed. Once the computer system is trained, it segregates the potential review population into responsive and non-responsive documents and prioritizes them in terms of relevancy (based on similarity to documents in the seed set) so that trained document reviewers can focus on documents that are likely to be most relevant first among the documents classified as responsive. Best practices also dictate that the producing party validate the results of its trained TAR system using certain metrics, such as a recall rate that measures the effectiveness of the software in finding responsive documents.

Following its production of documents from HPD and the Mayor's Office, the City commenced its review of ESI gathered from the DCP custodians and Banks (the "DCP/Banks review population"). At a conference on September 14, 2017, the City stated that it had "completed its discovery that was conducted in accordance with the court orders regarding [DCP] and Steve Banks." (Doc. No. 183 at 44:2-8.) When asked for further clarity, the City explained that it had applied search terms to the DCP/Banks review population that were targeted towards only a subset of Plaintiffs' Requests for Production that the Court had ordered

the City to search. It also indicated that it was unaware of whether Plaintiffs had been provided with the list of the search terms applied to the DCP/Banks review population. Accordingly, the Court directed the City to provide the search terms it used to Plaintiffs. The Court also stated that if Plaintiffs believed additional search terms should be applied to these custodians, Plaintiffs should propose the additional terms to the City within a week. The City then produced the list of search terms that were applied to the DCP/Banks review population, which was far more limited than the terms used for HPD and the Mayor's Office. Plaintiffs objected to the City's list of terms and proposed over 800 additional search terms to be run on the DCP/Banks review population. The City opposed Plaintiffs' request, contending that it would require the City to review approximately 90,000 additional documents. The City emphasized that it had thus far reviewed over 100,000 documents at a cost of over \$350,000, and that Plaintiffs' supplemental searches would result in additional costs of approximately \$248,000. (Doc. No. 203.)

After further negotiation and a modest reduction in Plaintiffs' number of search terms, the City agreed that it would accept all of Plaintiffs' modified proposed terms, but that it would use its predictive coding system to reduce the DCP/Banks review population to a more manageable number of documents. Plaintiffs, in turn, have indicated that the City's proposal is not acceptable because they are concerned about the reliability of the City's predictive coding processes. Specifically, Plaintiffs contend that the City has over-designated documents as privileged and non-responsive, including by applying an impermissibly narrow view of responsiveness during its review process. As a result of this alleged practice, Plaintiffs assert

that the City's TAR software was improperly trained on what constitutes a responsive and non-responsive document, leading to a predictive coding system that is unable to recognize documents that are truly responsive to the issues in this case.

In support of their allegations of over-designation, Plaintiffs produced to this Court certain documents that the City produced inadvertently, or produced in redacted format, which they contend should have been marked responsive and relevant. Among these were two electronic documents for which the City only produced a "slip sheet" because the documents had been designated as non-responsive (the "slip-sheeted" documents), but where Plaintiffs were nevertheless able to view the "extracted text" of the documents due to a production error. Plaintiffs contend that the slip-sheeted documents are in fact responsive to their document requests and, as a result, they should have been produced.

In response to Plaintiffs' concerns, this Court has issued several orders. It required the City to submit a letter for *in camera* review describing its predictive coding process and training for document reviewers. The City asserted that these materials are core work product, but nevertheless provided this information to the Court on September 8 and October 5, 2017. This Court also ordered briefing on Plaintiffs' challenges to the City's privilege designations on documents and in depositions. It required the City to provide a privilege log for a sample set of 80 documents that the City designated as privileged in its initial review. It also required submission of deposition transcripts and a privilege log concerning privilege objections made during the depositions.

In connection with the City's submission concerning its privilege assertions over the sample set of 80 documents, the City withdrew its privilege designation as to 36 documents and produced them. It also withdrew its privilege designation as to another 15 documents but it did not produce them because it asserted they were non-responsive. Out of the 80 document sample, the City maintained its original privilege assertion(s) over only 20 documents. This Court subsequently ordered the City to submit all 80 documents to this Court for *in camera* review and a more detailed log for purposes of assessing the validity of the remaining privilege designations.⁵ The City submitted the documents and detailed privilege log.

Finally, this Court ordered the City to provide an explanation for the allegedly non-responsive "slip-sheeted" documents referenced above, for which Plaintiffs could still view the extracted text. On November 3, 2017, the City provided its explanation and conceded that these slip-sheeted documents are "arguably responsive" to Plaintiffs' document requests. The City explained that the two documents had been designated "non-responsive," but since they were part of a partially responsive family of documents, the two documents were produced as a non-responsive "slip-sheet." The City further advised that these two documents had been reviewed as part of the City's "linear" review and were used in the City's initial TAR training round.

In response to the City's explanation about the "slip-sheeted documents," Plaintiffs filed a letter that attached three more examples of documents that had been "slip-sheeted" as non-

⁵ At this Court's direction, the City is re-reviewing its privilege designations to determine whether its privilege objections can be withdrawn as to any other documents. This Court will address Plaintiffs' challenge to the City's privilege designations in a separate order.

responsive, but that Plaintiffs nevertheless were able to view the extracted text of the documents. Plaintiffs contend that these three documents also should have been marked as responsive and produced.

This Court has now reviewed all of the parties' submissions with regard to the City's electronic document production and addresses Plaintiffs' current concerns below.

D. *Plaintiffs' Current Motion*

As referenced above, the instant dispute concerns three interrelated issues. Plaintiffs believe the City's TAR training and review process was faulty insofar as the City: (1) failed to use appropriate search terms for the DCP/Banks review population, resulting in a deficient potential review corpus from these 6 custodians; (2) searched for documents in the DCP/Banks review population that were responsive to only a subset of Plaintiffs' document requests; and (3) over-designated documents as non-responsive when in fact they were responsive to their document requests.

To increase transparency and confidence in the City's document review, Plaintiffs have proposed that the Court direct the City to provide for review by Plaintiffs' counsel random samples of each of the following (size and methodology to be negotiated by counsel and their e-discovery experts, or determined by the Court in the absence of agreement):

- non-privileged documents collected from DCP and Banks that were subject to TAR that are above the predictive-coding ranking cut-off used by the City, but where the City ultimately determined that the documents were non-responsive;
- non-privileged documents collected from DCP and Banks that were subject to TAR that are just below the predictive-coding ranking cut-off used by the City;

- non-privileged documents collected from HPD and the Mayor's Office that were subject to TAR that were above the predictive-coding cut-off, but that the City ultimately designated as non-responsive through its attorney review; and
- non-privileged documents collected from HPD and the Mayor's Office that were subject to TAR that were just below the predictive-coding cut-off used by the City.

Plaintiffs also seek:

- information about the ranking system used (*i.e.*, what cut-off was used, and how many documents were deemed responsive and unresponsive at each ranking);
- all materials submitted by the City for the Court's *in camera* review relating to predictive coding, other than materials that are in fact privileged, on an attorneys' and experts' eyes-only basis;
- an order requiring the City to search for documents responsive to all of Plaintiffs' document requests other than those specifically limited or stricken by the Court at prior court conferences (or those that had been withdrawn by Plaintiffs); and
- an order requiring the City to commence review of the additional search terms for the DCP/Banks review population document review process immediately (as opposed to waiting until after the settlement conference as the City proposes) and turning first to documents from Banks (so that some subset of his documents can be produced before his November 29 deposition).

In response, the City filed a letter motion requesting that the Court decline to consider Plaintiffs' application. (Doc. No. 209.)

DISCUSSION

I. LEGAL STANDARDS

Federal Rule of Civil Procedure 34 governs the production of documents and electronically stored information ("ESI"). In general, Rule 34 provides that a party may serve on any other party a request to produce documents and ESI within the party's possession, custody, or control so long as the documents and information are "relevant to any party's claim or defense and proportional to the needs of the case, considering the importance of the issues at stake in the action, the amount in controversy, the parties' relative access to relevant

information, the parties' resources, the importance of the discovery in resolving the issues, and whether the burden or expense of the proposed discovery outweighs its likely benefit." Fed. R. Civ. P. 34(a)(1) & 26(b)(1). Federal Rule of Civil Procedure 26(b)(1) also makes clear that discovery is not limited to information that will be admissible at trial.

Federal Rule of Civil Procedure 34(b)(2) specifies the obligations of a party who is on the receiving end of requests for documents and ESI. For each item or category of documents and information requested, the responding party must "state with specificity" the grounds, if any, for objecting to the request and, if there is an objection, "state whether any responsive materials are being withheld on the basis of that objection." Fed. R. Civ. P. 34(b)(2)(B)-(C).

Under Rule 26(g), attorneys for a party responding to requests for documents and information must sign discovery responses certifying, among other things, that to the best of the attorney's "knowledge, information, and belief formed after a reasonable inquiry," the document production is "complete and correct as of the time it is made," and that any objection or response is "not interposed for any improper purpose," including delay. Fed. R. Civ. P. 26(g)(1).

II. APPLICATION OF LEGAL STANDARDS TO THE INSTANT DISCOVERY DISPUTES

A. Search Terms To Be Used For The DCP/Banks Review Population

Collection, review, and production of ESI presents special challenges and requires "cooperation between opposing counsel and transparency in all aspects of preservation and production of ESI." *William A. Gross Const. Assocs., Inc. v. Am. Mfrs. Mut. Ins. Co.*, 256 F.R.D. 134, 136 (S.D.N.Y. 2009). Courts have recognized that keyword searches used to search for and

collect documents for review must be carefully crafted. *Id.* Here, Plaintiffs have provided 665 additional search terms to be applied to the DCP/Banks review population. The City has stated that the supplemental search would require review of 90,000 additional documents at a cost of approximately \$248,000. Nevertheless, it has stated that it is willing to use all of Plaintiffs' proposed search terms and use TAR, leveraging the training already done on the software. (Doc. No. 203.) Plaintiffs object to the City's proposal because they believe that the City's TAR processes are flawed insofar as it results in the over-designation of documents as non-responsive.

For the reasons set forth below in Section C, *infra*, this Court finds that there is no evidence of gross negligence or unreasonableness in the City's TAR training or review processes. Therefore, the basis for Plaintiffs' objections to the City's search term proposal appear to be moot. The Court accordingly orders the City to conduct a supplemental review of the DCP/Banks review population using the search terms proposed by Plaintiffs and leveraging the City's existing TAR processes.

B. Plaintiffs' Challenges To The City's Responsiveness Designations For The DCP/Banks Review Population

Next, the parties vigorously dispute what discovery must be produced from the DCP/Banks review population. This dispute stems from a Court conference held on February 16, 2017. On February 16, 2017, the Court ruled on the City's objections to 23 of Plaintiff's document requests and limited the documents and information that Plaintiffs could obtain in

response to this subset of Plaintiff's document requests.⁶ These limitations concerned both the topics of discovery and custodians (including DCP custodians and Banks) from whom discovery could be sought in connection with the 23 document requests. While Plaintiffs acknowledge that discovery responsive to the 23 document requests addressed during the conference has been curtailed by the Court's directives, they contend that the City is obligated to search for, and produce, documents from DCP and Banks that are responsive to the remaining 60 document requests tendered by Plaintiffs that were not discussed on February 16, 2017. The City, for its part, contends that it was only required to produce documents from DCP and Banks if the Court affirmatively ordered production during the February 16, 2017 conference.

To some extent, both sides misconstrue the Court's rulings during the February 16, 2017 conference. While the Court did grant discovery from DCP custodians and Banks as to certain the document requests that were most relevant to the issues in this case, nothing in the Court's ruling was intended to preclude—or grant—discovery of information responsive to Plaintiffs' other requests that were not specifically before the Court, provided that such information is relevant to the claims and defenses and proportional to the needs of the case.

As articulated during the February 16, 2017 conference, this Court has previously held that Plaintiffs' requests for documents and information from the DCP custodians and Banks must be related to the key issue in this litigation—the Community Preference Policy. By way of one example only, the Court ordered the City to search for and produce documents from DCP

⁶ The Court also addressed appropriate custodians during a conference on February 23, 2017, but that discussion was not tethered to any specific document requests.

custodians that discuss or consider the City’s federal “Affirmatively Furthering Fair Housing” obligations in the context of the Community Preference Policy. (*See* Doc. No. 87 at 36:21-38:21.) Similarly, with respect to Banks, the Court ruled that Plaintiffs were entitled to discovery regarding homeless individuals and the Community Preference Policy. (*See* Doc. No. 87 at 79:6-29.) These limitations appropriately narrowed Plaintiffs’ specific requests to information that is most relevant to this litigation. They also are proportional to the needs of the case because, as Plaintiffs acknowledge in their Complaint, neither DCP nor Banks are responsible for overseeing the Community Preference Policy, and the City already has produced discovery from HPD, the agency that is responsible for the Policy. Due to the number of custodians and volume of ESI at issue in this case, the City’s search for documents from DCP/Banks should be more limited than the search conducted of HPD or the Mayor’s Office because: (i) much of the relevant correspondence would likely involve HPD or the Mayor’s Office and, as a result, would be redundant; and (ii) there is no reason for the City to expend time and resources on the collection and review of ESI that concern functions of DCP and Banks that are unrelated to the Community Preference Policy and its role in overcoming community opposition to affordable housing. *See* Fed. R. Civ. P. 26(b)(2)(C) (courts are required to limit discovery that is unreasonably cumulative or duplicative).

In sum, the City must expand its search for documents responsive to Plaintiffs’ document requests as it construed this Court’s prior ruling too narrowly. If the City finds additional documents in the DCP/Banks review population regarding the primary issues and defenses in this case responsive to Plaintiffs’ other document requests—such as documents

concerning the Community Preference Policy and the City's justification for the same, including the Policy's role in overcoming community opposition—such documents should be produced, absent privilege or other objection for which the City has a good-faith basis to withhold the document. If, after review of the City's production of responsive documents from the DCP/Banks review population, Plaintiffs believe that the City impermissibly withheld documents responsive to specific requests, Plaintiffs shall meet-and-confer with the City and thereafter may file a motion to compel.

C. Plaintiffs' Challenges To The City's TAR Processes and Responsiveness Designations

Finally, Plaintiffs object to the City's continued use of its TAR system. Plaintiffs contend that the system is improperly trained because the City's human document reviewers over-designated documents as non-responsive during both the linear review and during the TAR training stages. As a result, Plaintiffs claim, the TAR software is unable to recognize and properly categorize responsive documents.

As courts have noted, the producing party is in the best position to "evaluate the procedures, methodologies, and technologies appropriate for preserving and producing their own electronically stored information." *Hyles v. New York City*, No. 10-cv-3119 (AT) (AJP), 2016 WL 4077114, at *3 (S.D.N.Y. Aug. 1, 2016) (citing Principle 6 of the Sedona Conference).

Traditionally, courts have not micro-managed parties' internal review processes for a number of reasons. First, attorneys, as officers of the court, are expected to comply with Rules 26 and 34 in connection with their search, collection, review and production of documents, including ESI. Second, internal attorney ESI work processes may reveal work product, litigation tactics,

and trial strategy. *See generally Disability Rights Council of Greater Wash. v. Wash. Metro. Transit Auth.*, 242 F.R.D. 139, 142-43 (D.D.C. 2007) (holding that a compilation of documents culled from a larger protection is protectable as attorney work product). Third, as noted above, the producing party is better equipped than the court to identify and utilize the best process for producing their own ESI consistent with their obligations under the Federal Rules of Civil Procedure. *See Hyles*, 2016 WL 4077114, at *3 (citing Principle 6 of the Sedona Conference). Fourth, perfection in ESI discovery is not required; rather, a producing party must take reasonable steps to identify and produce relevant documents. *See HM Elecs., Inc. v. R.F. Techs., Inc.*, No. 12-cv-2884 (BAS) (MDD), 2015 WL 471498, at *12 (S.D. Cal. Aug. 7, 2015), *vacated in part on other grounds*, 171 F. Supp. 3d 1020 (S.D. Cal. 2016); *Pension Comm. of the Univ. of Montreal Pension Plan v. Banc of Am. Secs., LLC*, 685 F. Supp. 2d 456, 461 (S.D.N.Y. 2010); *see also* Advisory Committee Notes, 2015 Amendments to Fed. R. Civ. P. 37(e). However, “[p]arties cannot be permitted to jeopardize the integrity of the discovery process by engaging in halfhearted and ineffective efforts to identify and produce relevant documents.” *HM Elecs., Inc.*, 2015 WL 4714908 at *12 (quoting *Bratka v. Anheuser-Busch Co., Inc.*, 164 F.R.D. 448, 463 (S.D. Ohio 1995)). “Litigation is not a game. It is the time-honored method of seeking the truth, finding the truth, and doing justice.” *Id.* (citation and quotations omitted). In keeping with these principles, this Court is of the view that there is nothing so exceptional about ESI production that should cause courts to insert themselves as super-managers of the parties’ internal review processes, including training of TAR software, or to permit discovery about such process, in the absence of evidence of good cause such as a showing of gross negligence in the review and

production process, the failure to produce relevant specific documents known to exist or that are likely to exist, or other malfeasance.

Courts are split as to the degree of transparency required by the producing party as to its predictive coding process. *See Rio Tinto PLC.*, 306 F.R.D. at 128 (citing John M. Facciola & Philip J. Favro, *Safeguarding the Seed Set: Why Seed Set Documents May Be Entitled To Work Product Protection*, 8 Fed. Cts. L. Rev. 1 (2015)). In some cases, parties have agreed to—or courts have ordered—transparency into the TAR or predictive coding processes. *See, e.g., id.* at 129 (“[t]he Court, however, need not rule on the need for seed set transparency in this case, because the parties agreed to a protocol that discloses all non-privileged documents in the control sets.”); *In re Actos Prods. Liab. Litig.*, No. 6:11-md-299, 2012 WL 7861249, at *4-5 (W.D. La. July 27, 2012) (the parties’ protocol had “experts” from each side simultaneously reviewing and coding the seed set); *Bridgestone Ams., Inc. v. Int’l Bus. Machs. Corp.*, No. 3:13-cv-1196, 2014 WL 4923014, at *1 (M.D. Tenn. July 22, 2014) (party offered to provide responsive and non-responsive seed set documents—an offer endorsed by the court). In other cases, courts have not required production of seed set documents. *See In re Biomet M2a Magnum Hip Implant Prods. Liab. Litig.*, No. 3:12-md-2391, 2013 WL 6405156, at *1-2 (N.D. Ind. Aug. 21, 2013) (refusing to direct party to produce seed set, but encouraging transparency); *Aurora Coop. Elev. Co. v. Aventine Renewable Energy-Aurora W., LLC*, No. 4:12-cv-230, 2015 WL 10550240, at *2 (D. Neb. Jan. 6, 2015) (encouraging the parties to work cooperatively in developing a TAR process, but observing that the Federal Rules of Civil Procedure do not require a party to disclose information that is not relevant to any party’s claim or defense).

While it is true that Plaintiffs here do not have clear insight into the City's predictive coding process and training, this Court has required the City to provide *in camera* submissions addressing these subjects. These *in camera* submissions reveal that the City appropriately trained and utilized its TAR system. The City's seed set included over 7,200 documents that were reviewed by the City's document review team and marked as responsive or non-responsive in order to train the system. Its seed set included randomly selected documents, as well as pre-coded example documents such as the pleadings, Plaintiffs' document requests, and other relevant, privileged and non-privileged documents. At the conclusion of the TAR training phase, which include five full training rounds, the City conducted a validation process.

Moreover, the City provided detailed training to its document review team as to the issues in the case. The document review team also was provided with all of Plaintiffs' document requests to use in connection with their review and designation of documents as responsive and non-responsive. In sum, the City's training and review processes and protocols present no basis for finding that the City engaged in gross negligence in connection with its ESI discovery – far from it.

Plaintiffs' objections to the City's use of TAR largely stems from their beliefs that the City has over-designated documents as non-responsive. In support of this claim, Plaintiffs rely on a few documents that were inadvertently produced by the City that they contend are responsive to their discovery requests and should have been produced. Plaintiffs also point to the fact that the City has flip-flopped on its designation of documents when challenged. For example, the City has admitted that the two "slip sheeted" documents initially marked as non-responsive are

in fact “arguably responsive.”⁷ Although the City has not provided an explanation for the additional three “slip-sheeted” documents that Plaintiffs have identified, these documents also appear to be at least “arguably responsive” to Plaintiffs’ discovery requests.

Additionally, Plaintiffs emphasize that the City’s document review has resulted in many documents being categorized as responsive, but ultimately withheld on privilege grounds. As mentioned above, the City has produced a sample of 80 of these supposedly privileged documents for *in camera* review. And, as noted above, the City has withdrawn its claim of deliberative process privilege as to 59 of the documents initially designated as privileged (out of the 80-document sample set), but also has changed its responsiveness designation to non-responsive and, as a result, has withheld these 15 documents as non-responsive to Plaintiffs’ document requests (the “NR Documents”). This Court has reviewed these 15 NR Documents *in camera* and determined that nearly all of them are responsive to Plaintiffs’ document requests to some degree, though most are only of marginal relevance.

In sum, the City incorrectly categorized at least five “slip-sheeted” documents as non-responsive during the electronic review process and correctly categorized 15 other documents as responsive during the electronic review process but later labeled them as non-responsive. However, this Court does not find the labeling of these 20 documents, only 5 of which were “incorrectly” categorized as non-responsive during the initial ESI review—out of the 100,000 documents that have been reviewed thus far in this case—sufficient to question the accuracy

⁷ Without ruling as to any objections on relevancy grounds, admissibility or privilege, this Court finds that in fact these documents are responsive to Plaintiff’s broad document requests.

and reliability of the City's TAR process as a whole. In any ESI review, "the Federal Rules of Civil Procedure do not require perfection." *Moore*, 287 F.R.D. at 191. Instead, the proper inquiry is whether the "search results are reasonable and proportional." *Hyles*, 2016 WL 4077114, at *3 (citing Fed. R. Civ. P. 26(g)(1)(B)).

Here, neither this Court nor Plaintiffs have identified anything in the TAR process itself that is inherently defective; rather, Plaintiffs' objections are premised upon human error in categorizing a small subset of documents as responsive or non-responsive. The City in this case has produced over 12,500 documents that were all designated as responsive. Moreover, it seems unlikely that the misdesignations identified by Plaintiffs would have affected the City's TAR processes in any meaningful way, given that the seed set was comprised of over 7,000 documents. The City's validation process, which was described to this Court in the City's *in camera* submission, further supports this conclusion.

While the Court disagrees with Plaintiffs' assertions that the TAR process as a whole is defective, it nevertheless finds that Plaintiffs have presented sufficient evidence to justify their request for sample sets of non-privileged documents from the documents pulled from the 50 custodians. In particular, this Court agrees that the sample sets will increase transparency, a request that is not unreasonable in light of the volume of documents collected from the custodians, the low responsiveness rate of documents pulled for review by the TAR software, and the examples that Plaintiffs have presented, which suggest there may have been some human error in categorization that may have led to gaps in the City's production.

Thus, this Court will grant Plaintiffs' request in part as follows:

- The City is directed to produce the five “slip-sheeted” documents and the 15 NR Documents within one week of the date of this order.
- The City is directed to provide to Plaintiffs a sample of 300 non-privileged documents in total from the HPD custodians and the Mayor’s Office.⁸ These documents should be randomly pulled from the corpus of non-responsive documents. The City shall provide these documents to Plaintiffs’ counsel on an attorneys’ and experts’ eyes-only basis. The City shall be required to produce the 300 non-privileged documents from the HPD custodians and the Mayor’s Office by December 22, 2017.
- The City shall apply the Plaintiffs’ search terms to the DCP/Banks review population and shall leverage its existing TAR process to segregate the additional documents into responsive and non-responsive categories and begin producing any responsive documents, subject to its objections and this Court’s orders limiting Plaintiffs’ discovery requests. It then shall provide Plaintiffs with a random sample of 100 non-privileged, non-responsive documents in total from the DCP/Banks review population.⁹ These too shall be provided to Plaintiffs’ counsel on an attorneys’ and experts’ eyes-only basis. The City shall be required to produce the 100 non-privileged documents from the DCP/Banks review population by January 12, 2018.
- To the extent Plaintiffs contend that the sampling process described above suggests that there are documents within the HPD/Mayor’s Office or DCP/Banks review populations that are responsive and relevant, but have not been produced, they shall meet and confer with the City to determine whether additional training and review is necessary with the understanding that reasonableness and proportionality, not perfection and scorched-earth, must be their guiding principles. *See Rio Tinto PLC*, 306 F.R.D. at 129 (“[o]ne point must be stressed—it is inappropriate to hold TAR to a higher standard than keywords or manual review. Doing so discourages parties from using TAR for fear of spending more in motion practice than the savings from using TAR for review.”).

⁸ The City may need to pull more than 300 documents to allow for the exclusion of any privileged documents that are pulled.

⁹ The City may need to pull more than 100 documents to allow for the exclusion of any privileged documents that are pulled.

This Court recognizes that the holiday season is upon us and that the application of over 600 search terms and sorting of documents for review purposes will most likely not be completed in advance of Banks' deposition on November 29. Nevertheless, this Court is of the view that Plaintiffs have sufficient information to depose Banks and therefore directs Plaintiffs to proceed with that deposition on November 29. To the extent there are privilege objections during his deposition, they should be marked for future ruling.

This Court denies Plaintiffs' request for information about the ranking system used by Plaintiffs (*i.e.*, what cut-off was used, and how many documents deemed responsive and unresponsive are at each ranking) and for all materials submitted by defendant *in camera* relating to predictive coding. Plaintiffs have failed to explain why information about the City's ranking system is needed, particularly given that Plaintiffs' request to sample non-responsive documents from the review populations is being granted. It is also unclear how this information is even potentially relevant to the claims and defenses in this litigation, as required under Federal Rule of Civil Procedure 26. However, in the interests of transparency and cooperation in the discovery process, the City is encouraged to share such information with Plaintiffs. With respect to the City's *in camera* submissions, this Court views this information as being protected by the work product privilege and, accordingly, is not subject to disclosure.

CONCLUSION

For the reasons set forth above, Plaintiffs' request (Doc. No. 208) is granted in part and denied in part. The City's request for the Court to disregard Plaintiffs' application (Doc. No. 209)

is denied as moot. The Clerk of Court is respectfully directed to terminate the gavel pending in connection with Doc. No. 209.

SO ORDERED.

Dated: November 27, 2017
New York, New York

A handwritten signature in black ink, reading "Katharine H. Parker". The signature is written in a cursive, flowing style. The first name "Katharine" is written in a larger, more prominent script, followed by "H." and "Parker". The signature is positioned above a horizontal line.

KATHARINE H. PARKER
United States Magistrate Judge